# METHOD AND APPARATUS FOR PROVIDING A VIRTUAL AGE ESTIMATION FOR REMAINING LIFETIME PREDICTION OF A SYSTEM USING NEURAL NETWORKS

5    Reference is hereby made to copending:

U.S. Provisional Patent Application No. 60/255,615 filed 12/14/2000 for NEURAL NETWORK-BASED VIRTUAL AGE ESTIMIATION FOR REMAINING LIFETIME, in the names of Christian Darken and Markus Loecher, Attorney Docket No. 00P9072US;

10    U.S. Provisional Patent Application No. 60/255,614 filed 12/14/2000 for POLYNOMIAL BASED VIRTUAL AGE ESTIIMATION FOR REMAINING LIFETIME PREDICTION, in the names of Markus Loecher and Christian Darken, Attorney Docket No. 00P9073US; and

U.S. Provisional Patent Application No. 60/255,613 filed 12/14/2000 for MARKOV

15    TRANSITION PROBABILITIES FOR PREDICTIVE MAINTENANCE, in the name of Markus Loecher, Attorney Docket No. 00P9074US,

of which priority is claimed and whereof the disclosures are hereby incorporated herein by reference.

Reference is also made to copending patent applications being filed on even date

20    herewith:

METHOD AND APPARATUS FOR PROVIDING A POLYNOMIAL BASED VIRTUAL AGE ESTIMATION FOR REMAINING LIFETIME PREDICTION OF A SYSTEM, in the names of Markus Loecher and Christian Darken, Attorney Docket No. 00P9073US01; and METHOD AND APPARATUS FOR PROVIDING PREDICTIVE

25    MAINTENANCE OF A DEVICE BY USING MARKOV TRANSITION PROBABILITIES, in the name of Markus Loecher, Attorney Docket No. 00P9074US01, and whereof the disclosures are hereby incorporated herein by reference.

The present invention relates generally to the field of failure prediction and, more specifically to deriving an estimate of the remaining lifetime of a generic system or apparatus.

Devices and apparatus used in various fields of medicine, industry, transportation,
5 communications, and so forth, typically have a certain useful or operational life, after which replacement, repair, or maintenance is needed. Generally, the expected length of the operational life is known only approximately and, not untypically, premature failure is a possibility. Simple running time criteria are typically inadequate to provide timely indication of an incipient failure. In some applications, unanticipated failure of devices
10 constitutes a at least a nuisance; however, more typically, unanticipated device failure may be a major nuisance leading to costly interruption of services and production. In other cases, such unexpected failure can seriously compromise safety and may result in potentially dangerous and life-threatening situations.

In accordance with an aspect of the invention, a complex function of monitored variables
15 is estimated and then used to compute its "virtual age", which is then compared with a fixed threshold.

In accordance with an aspect of the invention, an approach is utilized for the general task of failure prediction, which is part of a condition based or predictive maintenance.

In accordance with an aspect of the invention, a method of virtual age estimation for
20 remaining lifetime prediction incrementally augments a "virtual age" by continuously monitoring significant parameters of a system throughout at least a portion of its active life.

In accordance with an aspect of the invention, the functional form of the state-dependent virtual age or wear increment is taken to be a radial basis function (RBF) neural network
25 whereof the coefficients are obtained in a training phase.

In accordance with an aspect of the invention, a method for providing a virtual age estimation for predicting the remaining lifetime of a device of a given type, comprises the steps of monitoring a predetermined number of significant parameters of respective ones

of a training set of devices of the given type, the parameters contributing respective wear increments, determining coefficients of a radial basis function neural network for modeling the wear increments determined from the training set operated to failure and whereof the respective virtual ages are normalized substantially to a desired norm value,

5   deriving from the radial basis function neural network a formula for virtual age of a device of the given type, and applying the formula to the significant parameters from a further device of the given type for deriving wear increments for the further device.

The method will be more fully understood from the following detailed description of preferred embodiments, in conjunction with the Drawing, in which

10  Figure 1 shows a diagrammatic flow chart of steps in accordance with the principles of the invention; and

Figure 2 shows a block diagram for apparatus in accordance with the principles of the invention.

In Figure 1, step 2 involves collecting data histories of devices until failure. In general

15  this will conform to a matrix with N rows (uses) and M columns (variables).

In step 4 a clustering algorithm is applied to partition the data set into Z clusters. The centers and widths of Gaussian radial basis functions are fixed.

In step 6 the data matrix C is computed, solving for linear weights $a$ using Ridge regression. Cross validation is used to optimize.

20  In step 8, linear weights $a$ and cluster centers and widths are used to compute wear increments for devices in operation.

In step 10, the sum of wear increments, that is, the virtual age, is compared with a user specified threshold and if the threshold is exceeded, a warning indication or signal is given.

25  12 generally indicates the use of cross validation to optimize the number of variables M to be used and the number of clusters.

As shown in Figure 2, a computer 20 is equipped with data and program storage equipment 22 and a source 26 of programs for training and operating in an interactive manner as hereinafter described. Data from training sessions as further explained below is provided at 24. A device or system 28 which is being monitored provides data by way of data collection interface unit 30 to computer 20. Computer 20 provides an imminent or prospective alarm as to lifetime expiration and/or failure expectation at an alarm device 32.

The method in accordance with the present invention is widely applicable in many fields. In order to facilitate understanding of the invention and to illustrate the use of device-specific information and parameters, the invention will next be more fully described by way of an exemplary, non-limiting embodiment relating to X-ray tubes; where appropriate, generally applicable notions are also stated in the context of the specific exemplary embodiment. The example used is also appropriate in that an unexpected failure of such an X-ray tube, for example during a critical surgical procedure, should be avoided insofar as is possible.

Suppose, $x_n = (x_{1,n} \ldots x_{d,n})$ is a time-series of d-dimensional measurement vectors. The individual scalars $x_i$ could be any quantity affecting the rate of wear or ageing of the device, including directly measured physical quantities such as temperature or voltage or composite functions thereof such as, for example, power (product of voltage and current) or temperature difference, or e.g. control parameters such as load settings and time of operation. The choice of both the number d and kind of variables, which usually is only a small subset of available measurements, can be done following existing variable selection techniques. In the X-ray tube case, it turns out to have been possible to perform an exhaustive search, which selected the d unique scalars that minimized the cross validation (CV) error as will be explained in more detail below.

During the life of the device there will be typically many thousands of vectors, each of which contributes a small increment to the total wear. Without loss of generality, it is herein proposed to reduce the uncertainty in remaining lifetime estimation by the following method:

4

The wear increment f() is modeled by a radial basis function neural network with M hidden units:

$$f(\vec{x}_n) = \sum_{i=0}^{M-1} a_i g(\vec{x}_n, \vec{z}_i, \sigma_i) \qquad (1)$$

, where g is a radially-symmetric function centered at $z_i$ with width parameter $\sigma_i$. The

5   number of units M is a free parameter, which again should be optimized by cross validation.

In the case of the X-ray tube, this form was found to be optimal. In general, the normalized form

$$f(\vec{x}_n) = \sum_{i=1}^{M} a_i g(\vec{x}_n, \vec{z}_i, \sigma_i) \bigg/ \sum_{j=1}^{M} g(\vec{x}_n, \vec{z}_j, \sigma_j)$$

10  may be used. In either case, the weights $a_i$ enter this equation linearly and hence can be solved for using linear methods, whereas the internal parameters $z_i$ and $\sigma_i$ must be obtained through nonlinear techniques.

For the case of Gaussian basis function, which was found to be appropriate and was used for the X-ray tubes, we have

15
$$g(\vec{x}, \vec{z}, \sigma) = \exp\left(-\frac{\|\vec{x} - \vec{z}\|^2}{2\sigma^2}\right)$$

The $z_i$ can be selected by applying a clustering algorithm, such as k-means, to the measurement vectors. The $\sigma_i$ can be selected in one of several ways, e.g.

- $\sigma_i$ can be taken to be the distance from the i'th measurement to the first (or k'th) nearest measurements. This method was chosen for the X-ray tubes.

20  - $\sigma_i$ can be taken to be a global constant, e.g. the average of the distance from each measurement to the first (or k'th) nearest measurement.

- In either of the above cases, a scaling factor can be applied. This would introduce another free parameter $\lambda$ ($\sigma_i$ transforms into $\lambda\sigma_i$) to be chosen via cross-validation.

Note that equation (1) can be conveniently rewritten into a sum of $M$ terms of the form

$$f(\vec{x}_n) = \sum_{j=0}^{M-1} a_j f_j(\vec{x}_n) \quad (2)$$

, where M is the number of coefficients $a_j$. The dependence on the $z_i$ and the $\sigma_i$ is hidden, as these parameters are fixed through the methods described above. Now we are left with a linear system of equations. We determine the M coefficients $a_j$ in the supervised training phase as follows:

Suppose, there are N device histories of tubes, which eventually failed, indexed by $k$. This constitutes our training set. Denote the number of vectors for each device by $L_k$. For each device we compute the M independent sums over all wear increments, which naturally depend on the M unknown coefficients:

$$C_{k,j} = \sum_{n=1}^{L_k} f_j(\vec{x}_n^k)$$

This yields a (N x M) matrix $(C)_{k,j}$ and $N$ equations for the virtual age of each device, which have the form

$$(VirtualAge)_k = \sum_{j=0}^{M-1} a_j C_{k,j}$$

Ideally, the virtual ages for each device would be identical, say one. In order to find the best weights, such that all virtual ages are as closes as possible to an arbitrary constant (we choose 1), we propose to minimize the sum-of-squared-error criterion function

$$J(\vec{a}) = \left\| \overline{\overline{C}} \cdot \vec{a} - \vec{1} \right\|^2 + \lambda \vec{a}^T \overline{\overline{B}} \vec{a}$$

The first term on the right side corresponds to the ordinary linear least squares regression. The additional term involving $\lambda$, is intended to improve the generalizability and avoid over fitting. This technique is referred to as ridge regression in the pertinent literature. The parameter $\lambda$ should be optimized via cross validation. The matrix $B$ is positive definite and for the X-ray tubes was taken to be the identity matrix.

In the case of missing data, i.e. if for a particular device $z$ only a fraction $f_k$ of data is available, we have to replace the constant vector 1 with the device dependent vector f:

$$J(\vec{a}) = \left\| \overline{\overline{C}} \cdot \vec{a} - \vec{f} \right\|^2 + \lambda \vec{a}^T \overline{\overline{B}} \vec{a}$$

After determining the coefficients **a** for the N devices in the training set, it is proposed in accordance with the embodiment of the invention to estimate the remaining lifetime of devices in the same family by computing the incremental (and resulting cumulative) wear according to equation (2). Since the virtual age was normalized to one (1), the cumulative wear directly yields the fractional life that has elapsed.

The applicability of the principles of cross correlation in the context of the present invention is next addressed. K-fold cross validation is a well known technique to estimate the test error of a predictor if the available data set (size $n$) is too small to allow the split into training and test sets. Instead, we iterate the splitting process by dividing the data into a "small" part of $k$ elements and use the remaining $n-k$ elements for training. The test errors on the small k-set are then averaged to yield the k-fold cross validation error. In the X-ray tube example, the data set comprised approximately 70 tubes (n~70) and we chose k ~ 1-5.

It will be understood that the invention may be implemented in a number of ways, utilizing available hardware and software technologies. Implementation by way of a programmable digital computer is suitable, with or without the addition of supplemental apparatus. A dedicated system may also be used, with a dedicated programmed computer and appropriate peripheral equipment. When various functions and subfunctions are implemented in software, subsequent changes and improvements to the operation are readily implemented.

While the present invention has been described by way of illustrative embodiments, it will be understood by one of skill in the art to which it pertains that various changes and modifications may be made without departing from the spirit of the invention. Such changes and modifications are intended to be within the scope of the claims following.